S21-47

.127521

N94-22314

Automated Cloud Classification With A Fuzzy Logic Expert System

Vasanth Tovinkere[1] and Bryan A. Baum[2]

[1]Lockheed Engineering and Sciences Co., Hampton, VA 23666
[2]NASA Langley Research Center, Hampton, VA 23681-0001

1. Introduction

An unresolved problem in current cloud retrieval algorithms concerns the analysis of scenes containing overlapping cloud layers. Cloud parameterizations are very important both in global climate models and in studies of the Earth's radiation budget. Most cloud retrieval schemes, such as the bispectral method used by the International Satellite Cloud Climatology Project (ISCCP), have no way of determining whether overlapping cloud layers exist in any group of satellite pixels. One promising method uses fuzzy logic to determine whether mixed cloud and/or surface types exist within a group of pixels, such as cirrus, land, and water, or cirrus and stratus. When two or more class types are present, fuzzy logic uses membership values to assign the group of pixels partially to the different class types. The strength of fuzzy logic lies in its ability to work with patterns that may include more than one class, facilitating greater information extraction from satellite radiometric data. The development of the fuzzy logic rule-based expert system involves training the fuzzy classifier with spectral and textural features calculated from accurately labelled 32x32 regions of Advanced Very High Resolution Radiometer (AVHRR) 1.1-km data. The spectral data consists of AVHRR channels 1 (0.55-0.68 μm), 2 (0.725-1.1 μm), 3 (3.55-3.93 μm), 4 (10.5-11.5 μm), and 5 (11.5-12.5 μm), which include visible, near-infrared, and infrared window regions. The textural features are based on the gray level difference vector (GLDV) method. A sophisticated new interactive visual image classification system (IVICS) (Tovinkere et al., 1993) is used to label samples chosen from scenes collected during the FIRE IFO II. The training samples are chosen from predefined classes, chosen to be ocean, land, unbroken stratiform, broken stratiform, and cirrus. The November 28, 1991 NOAA overpasses contain complex multilevel cloud situations ideal for training and validating the fuzzy logic expert system.

2. Classification scheme

(i) *Features and Classes*

Both textural and spectral features were computed for each of the regions for use as features. The textural features were computed using the GLDV approach (Haralick et al, 1973; Weszka et al, 1976; Chen et al. 1989). As noted by Tovinkere et al. (1993), the set of features used by Ebert (1987), Garand (1988), and Tovinkere et al. (1993) may not necessarily compose the optimum feature vector, i.e., the best choice of features. There are alternative feature sets that may perform as well or better. For this investigation, the feature vector was determined through a sequential forward selection (SFS) procedure (Devijer and Kittler 1982), which is a simple bottom-up procedure where one measurement at a time is added to the current feature set. At each stage, the attribute to be included in the feature set is selected from the remaining available measurements, so that the new set yields a maximum value of a criterion function used. The feature selector uses the Jeffries-Matusita distance (also known as the Bhattacharya distance) as the criterion function (Welch et. al., 1992). A second level of feature selection is performed on the feature set selected by the SFS algorithm. This eliminates redundant information and comes up a minimal set required to discriminate the given classes. The measures chosen for this investigation have both textural and spectral features. The spectral features are (1) mean gray level (GL) difference of channels 2 and 3 (GL 2-GL 3) and (2) mean brightness temperature of channel 4 (Rmean 4). The textural features include entropy (ENT) of channels 1 and 3, homogeneity (HOM) of channel 3 and the angular second moment (ASM) of channel 1. Details for computing these textural measures can be found in Welch et al. (1992). Each of these features uniquely describes at least one class and are chosen for daytime classification analysis.

The spectral feature GL 2-GL 3 is computed by scaling the reflectance of channel 2 to lie between 0 and 255. The range of values between 0- 255 represent the reflectivities between 0%-100%. Similarly, channel 3 brightness temperature is scaled to lie between 0-255 which represents temperatures between 200-327.5 K in half degree increments. The mean gray level for channels 2 (GL 2) and channels 3 (GL 3) are obtained by computing the mean of the gray levels of all pixels in the 32x32 region for the respective channels. The feature Rmean 4 is mean brightness temperature of channel 4 for the 32x32 pixel region.

A number of scenes from the IFO were used to derive training and testing samples. The mean (μ) and the standard deviations (σ) are calculated for the complete training data set. The graphs in Figure 1 show the extent of each feature for each class, i.e., ($\mu - \sigma$) to ($\mu + \sigma$). These graphs depict the separability of the classes for each feature. The data from the NOAA-11 and NOAA-12 overpasses on November 28, 1991, are classified into the following classes: (1) water, (2) land, (3) unbroken stratiform,(4) broken stratiform, and (5) cirroform. These classes are broad in scope and may contain a number of representative subclasses. For instance, land covers all surface not covered by water, unbroken stratiform includes both stratus and altostratus cloud types, broken stratiform includes both stratus and altostratus cloud types in which some surface is visible in the 32 by 32 pixel sample, and cirroform includes cirrostratus, cirrus uncinus, and other cirrus types.

(ii) Training of the fuzzy classifier

The fuzzy logic classifier is trained and tested using the 1.1-km AVHRR data. A number of scenes from the FIRE IFO II time period were used to derive numerous 32 by 32 pixel subregions of "pure" classes for classification purposes. Region labeling was performed using the IVICS system (Tovinkere et al., 1993). Details of the fuzzy logic-based expert system can be found in Tovinkere et al. (1993). The number of classes is limited for this initial study to ease interpretation of the classification results for mixed cloud and/or surface types.

3. Results

(i) Classification of pure samples

Two thirds of the data set from the spectral and textural database, with replacement, is used to train the classifier and the remainder is used as the test data to determine the accuracy. Replacement of the sample allows each sample to be selected more than once to generate the training data. This insures an unbiased estimate of classification accuracy. Table 1 shows the misclassification chart for the classifier.

Table 1: Misclassification chart for the fuzzy classifier.

CLASSES	1	2	3	4	5	Percentage Accuracy
Ocean/Water	57	2	0	0	0	96.61
Land	2	62	0	0	0	96.88
Unbroken Stratiform	0	0	71	2	0	97.26
Broken Stratiform	0	0	3	80	0	96.39
Cirroform	0	0	1	0	130	99.24
Overall Accuracy						**97.56**

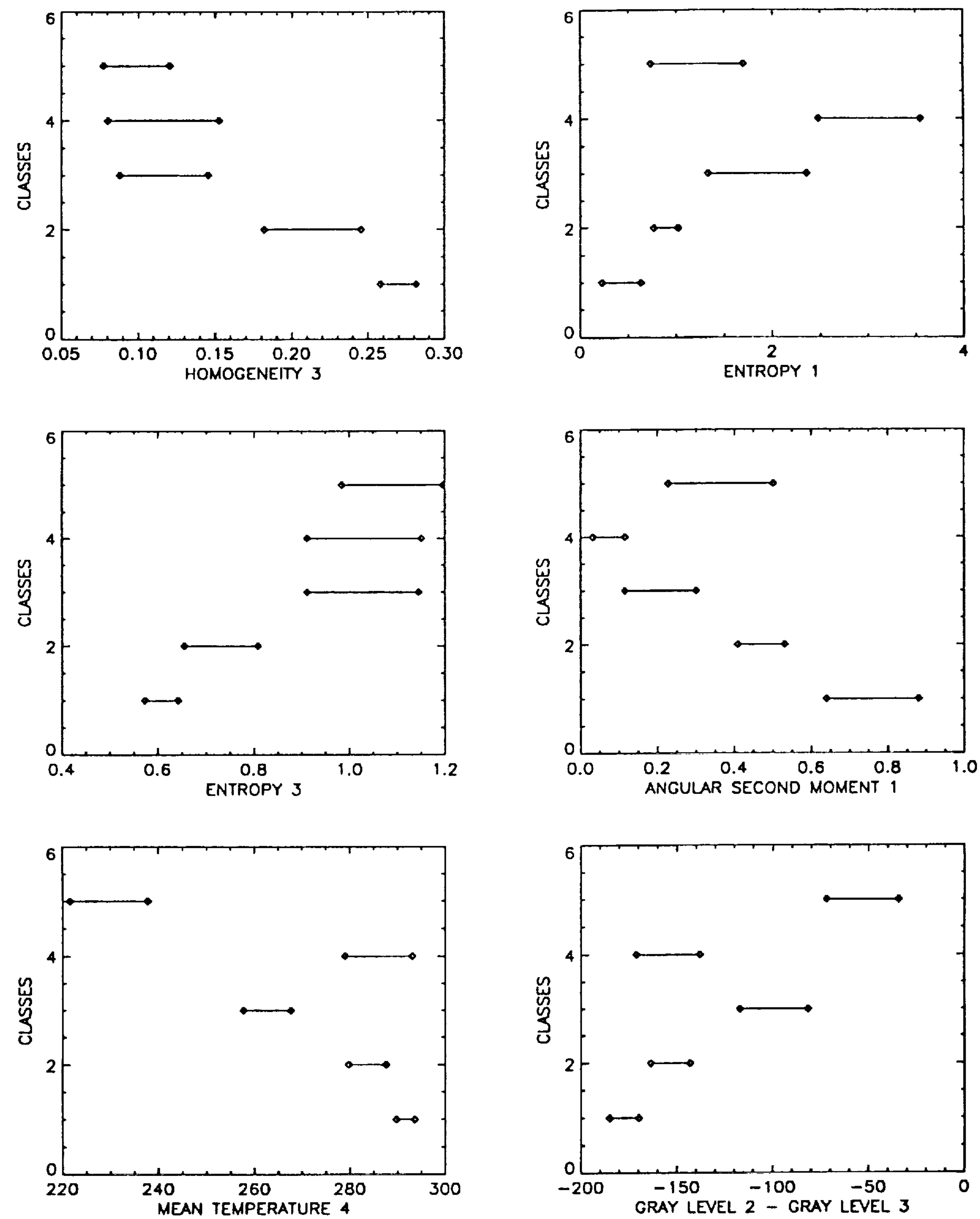

Figure 1.–Class means and standard deviations for all the features used in the fuzzy classifier. The classes considered in the above case are: 1. Ocean/Water, 2. Land, 3. Unbroken Stratiform, 4. Broken Stratiform, and 5. Cirroform. The feature set under consideration is for daytime analysis only as they rely on the visible channels.

(ii) Classification of mixed scenes

The classifier was tested for 100-200 samples consisting of multilayer clouds. The classifier was able to determine the various classes present in regions of patchy cirrus and when thin cirrus completely covered the stratiform clouds indicating a multilayer situation. However, it failed when optically thick cirroform clouds enveloped the region. The classifier was also tested on regions that contained mixed surface types. In most cases the classifier was able to identify the different surface types present in the region. More detailed results will be presented regarding the ability of the classifier to identify regions of mixed cloud types.

REFERENCES

Chen, D. W., S. K. Sengupta, and R. M. Welch, Cloud field classification based upon high spatial resolution textural features, Part 2, Simplified vector approaches, *J. Geophys. Res.*, *94*, 14,749-14,765, 1989.

Devijer, P. A., and J. Kittler, *Pattern Recognition: A Statistical Approach*, 448 pp, Prentice Hall, Englewood Cliffs, N. J., 1982.

Ebert, E., A pattern recognition technique for distinguishing surface and cloud types in polar regions, *J. Clim. Appl. Meterol.*, 26,1412-1427, 1987.

Garand, L, 1988: Automated recognition of coceanic cloud patterns. Part I: Methodology and application to cloud climatology. *J. Clim.*, 1, 20-39.

Giarratano, J., and G. Riley, *Expert Systems*, 632 pp., PWS-Kent, Boston, Mass., 1990.

Haralick, R., M., K. S. Shanmugam, and I. Dinstein, Textural features for image classification, *IEEE trans. Syst., Man Cybern.*, *SMC-3(6)*, 610-621, 1973.

Tovinkere, V. R., M. Penaloza, A. Logar, J. Lee, R. C. Weger, T. A. Berendes, and R. M. Welch, An intercomparison of artificial intelligence approaches for polar scene identification, *J. Geophys. Res.*, 98, 5001-5016, 1993.

Welch, R. M., S. K. Sengupta, A. K. Goroch, P. Rabindra, N. Rangaraj, and M. S. Navar, Polar cloud and surface classification using AVHRR imagery: An intercomparison of methods, *J. Appl. Meterol.*, *31*, 405-420, 1992.

Weszka, J. S., C. R. Dyer, and A. Rosenfeld, A comparative study of texture measures for terrain classification, *IEEE Trans. Syst. Man. Cybern.*, *SMC-6(4)*, 2269-2285, 1976.